

PCT

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/09/720,841
 DATE: 02/05/2001
 TIME: 16:34:36

Input Set : A:\5787585.txt
 Output Set: N:\CRF3\02052001\I720841.raw

ENTERED

```

3 <110> APPLICANT: Biotica Technology Limited
4   Leadlay, Peter F
5   Pfizer, Inc.
6   Staunton, James
7   Cortes, Jesus
8   McArthur, Hamish AI
10 <120> TITLE OF INVENTION: Polyketides and their synthesis
12 <130> FILE REFERENCE: IS/CP5787585
C--> 14 <140> CURRENT APPLICATION NUMBER: US/09/720,841
C--> 15 <141> CURRENT FILING DATE: 2000-12-29
17 <150> PRIOR APPLICATION NUMBER: GB 9814006.4
18 <151> PRIOR FILING DATE: 1998-06-29
20 <160> NUMBER OF SEQ ID NOS: 53
22 <170> SOFTWARE: PatentIn Ver. 2.1
24 <210> SEQ ID NO: 1
25 <211> LENGTH: 398
26 <212> TYPE: PRT
27 <213> ORGANISM: Streptomyces sp. C5
29 <400> SEQUENCE: 1
30 Met Val Thr Gly Leu Gly Ile Val Ala Pro Asn Gly Leu Gly Val Gly
31   1           5           10           15
33 Ala Ile Trp Asp Ala Val Leu Asn Gly Arg Asn Gly Ile Gly Pro Leu
34   20           25           30
36 Arg Arg Phe Ala Asp Asp Gly Arg Leu Gly Arg Leu Ala Gly Glu Val
37   35           40           45
39 Ser Asp Phe Val Pro Glu Asp His Leu Pro Lys Arg Leu Leu Val Gln
40   50           55           60
42 Thr Asp Pro Met Thr Gln Met Thr Ala Leu Ala Ala Ala Glu Trp Ala
43   65           70           75           80
45 Leu Arg Glu Ala Gly Cys Ala Pro Ser Ser Pro Leu Glu Ala Gly Val
46   85           90           95
48 Ile Thr Ala Ser Ala Ser Gly Gly Phe Ala Ser Gly Gln Arg Glu Leu
49   100          105          110
51 Gln Asn Leu Trp Ser Lys Gly Pro Ala His Val Ser Ala Tyr Met Ser
52   115          120          125
54 Phe Ala Trp Phe Tyr Ala Val Asn Thr Gly Gln Ile Ala Ile Arg His
55   130          135          140
57 Asp Leu Arg Gly Pro Val Gly Val Val Val Ala Glu Gln Ala Gly Gly
58 145          150          155          160
60 Leu Asp Ala Leu Ala His Ala Arg Arg Lys Val Arg Gly Gly Ala Glu
61   165          170          175
63 Leu Ile Val Ser Gly Ala Met Asp Ser Ser Leu Cys Pro Tyr Gly Met
64   180          185          190
66 Ala Ala Gln Val Arg Ser Gly Arg Leu Ser Gly Ser Asp Asp Pro Thr
67   195          200          205
69 Ala Gly Tyr Leu Pro Phe Asp Arg Arg Ala Ala Gly His Val Pro Gly
70   210          215          220

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/720,841

DATE: 02/05/2001

TIME: 16:34:36

Input Set : A:\5787585.txt

Output Set: N:\CRF3\02052001\I720841.raw

```

72 Glu Gly Gly Ala Ile Leu Ala Val Glu Asp Ala Glu Arg Val Ala Glu
73 225 230 235 240
75 Arg Gly Gly Lys Val Tyr Gly Ser Ile Ala Gly Thr Ala Ser Phe Asp
76 245 250 255
78 Pro Pro Pro Gly Ser Gly Arg Pro Ser Ala Leu Ala Arg Ala Val Glu
79 260 265 270
81 Thr Ala Leu Ala Asp Ala Gly Leu Asp Arg Ser Asp Ile Ala Val Val
82 275 280 285
84 Phe Ala Asp Gly Ala Ala Val Gly Glu Leu Asp Val Ala Glu Ala Glu
85 290 295 300
87 Ala Leu Ala Ser Val Phe Gly Pro His Arg Val Pro Val Thr Val Pro
88 305 310 315 320
90 Lys Thr Leu Thr Gly Arg Leu Tyr Ser Gly Ala Gly Pro Leu Asp Val
91 325 330 335
93 Ala Thr Gly Leu Leu Ala Leu Arg Asp Glu Val Val Pro Ala Thr Gly
94 340 345 350
96 His Val His Pro Asp Pro Asp Leu Pro Leu Asp Val Val Thr Gly Arg
97 355 360 365
99 Pro Arg Ala Met Ala Asp Ala Arg Ala Ala Leu Val Val Ala Arg Gly
100 370 375 380
102 His Gly Gly Phe Asn Ser Ala Leu Val Val Arg Gly Ala Ala
103 385 390 395
107 <210> SEQ ID NO: 2
108 <211> LENGTH: 425
109 <212> TYPE: PRT
110 <213> ORGANISM: Streptomyces peucetius
112 <400> SEQUENCE: 2
113 Met Thr Gly Thr Ala Ala Arg Thr Ala Ser Ser Gln Leu His Ala Ser
114 1 5 10 15
116 Pro Ala Gly Arg Arg Gly Leu Arg Gly Arg Ala Val Val Thr Gly Leu
117 20 25 30
119 Gly Ile Val Ala Pro Asn Gly Leu Gly Val Gly Ala Tyr Trp Asp Ala
120 35 40 45
122 Val Leu Asn Gly Arg Asn Gly Ile Gly Pro Leu Arg Arg Phe Thr Gly
123 50 55 60
125 Asp Gly Arg Leu Gly Arg Leu Ala Gly Glu Val Ser Asp Phe Val Pro
126 65 70 75 80
128 Glu Asp His Leu Pro Lys Arg Leu Leu Ala Gln Thr Asp Pro Met Thr
129 85 90 95
131 Gln Tyr Ala Leu Ala Ala Ala Glu Trp Ala Leu Arg Glu Ser Gly Cys
132 100 105 110
134 Ser Pro Ser Ser Pro Leu Glu Ala Gly Val Ile Thr Ala Ser Ala Ser
135 115 120 125
137 Gly Gly Phe Ala Phe Gly Gln Arg Glu Leu Gln Asn Leu Trp Ser Lys
138 130 135 140
140 Gly Pro Ala His Val Ser Ala Tyr Met Ser Phe Ala Trp Phe Tyr Ala
141 145 150 155 160
143 Val Asn Thr Gly Gln Ile Ala Ile Arg His Asp Leu Arg Gly Pro Val
144 165 170 175

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/720,841

DATE: 02/05/2001

TIME: 16:34:36

Input Set : A:\5787585.txt

Output Set: N:\CRF3\02052001\I720841.raw

```

146 Gly Val Val Val Ala Glu Gln Ala Gly Gly Leu Asp Ala Leu Ala His
147      180      185
149 Ala Arg Arg Lys Val Arg Gly Gly Ala Glu Leu Ile Val Ser Gly Ala
150      195      200      205
152 Val Asp Ser Ser Leu Cys Pro Tyr Gly Met Ala Ala Gln Val Lys Ser
153      210      215      220
155 Gly Arg Leu Ser Gly Ser Asp Asn Pro Thr Ala Gly Tyr Leu Pro Phe
156 225      230      235      240
158 Asp Arg Arg Ala Ala Gly His Val Pro Gly Glu Gly Gly Ala Ile Leu
159      245      250      255
161 Thr Val Glu Asp Ala Glu Arg Ala Ala Glu Arg Gly Ala Lys Val Tyr
162      260      265      270
164 Gly Ser Ile Ala Gly Tyr Gly Ala Ser Phe Asp Pro Pro Pro Gly Ser
165      275      280      285
167 Gly Arg Pro Ser Ala Leu Ala Arg Ala Val Glu Thr Ala Leu Ala Asp
168      290      295      300
170 Ala Gly Leu Asp Gly Ser Asp Ile Ala Val Val Phe Ala Asp Gly Ala
171 305      310      315      320
173 Ala Val Pro Glu Leu Asp Ala Ala Glu Ala Glu Ala Leu Ala Ser Val
174      325      330      335
176 Phe Gly Pro Arg Arg Val Pro Val Thr Val Pro Lys Thr Leu Thr Gly
177      340      345      350
179 Arg Leu Tyr Ser Gly Ala Gly Pro Leu Asp Val Ala Thr Ala Leu Leu
180      355      360      365
182 Ala Leu Arg Asp Glu Val Val Pro Ala Thr Ala His Val Asp Pro Asp
183      370      375      380
185 Pro Asp Leu Pro Leu Asp Val Val Thr Gly Arg Pro Arg Ser Leu Ala
186 385      390      395      400
188 Asp Ala Arg Ala Ala Leu Leu Val Ala Arg Gly Tyr Gly Gly Phe Asn
189      405      410      415
191 Ser Ala Leu Val Val Arg Gly Ala Ala
192      420      425
196 <210> SEQ ID NO: 3
197 <211> LENGTH: 407
198 <212> TYPE: PRT
199 <213> ORGANISM: Streptomyces coelicolor
201 <400> SEQUENCE: 3
202 Met Ser Val Leu Ile Thr Gly Val Gly Val Val Ala Pro Asn Gly Leu
203 1      5      10      15
205 Gly Leu Ala Pro Tyr Trp Ser Ala Val Leu Asp Gly Arg His Gly Leu
206      20      25      30
208 Gly Pro Val Thr Arg Phe Asp Val Ser Arg Tyr Pro Ala Thr Leu Ala
209      35      40      45
211 Gly Gln Ile Asp Asp Phe His Ala Pro Asp His Ile Pro Gly Arg Leu
212      50      55      60
214 Leu Pro Gln Thr Asp Pro Ser Thr Arg Leu Ala Leu Thr Ala Ala Asp
215      65      70      75      80
217 Trp Ala Leu Gln Asp Ala Lys Ala Asp Pro Glu Ser Leu Thr Asp Tyr
218      85      90      95

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/720,841

DATE: 02/05/2001

TIME: 16:34:36

Input Set : A:\5787585.txt

Output Set: N:\CRF3\02052001\I720841.raw

```

220 Asp Met Gly Val Val Thr Ala Asn Ala Cys Gly Gly Phe Asp Phe Thr
221      100      105      110
223 His Arg Glu Phe Arg Lys Leu Trp Ser Glu Gly Pro Lys Ser Val Ser
224      115      120      125
226 Val Tyr Glu Ser Phe Ala Trp Phe Tyr Ala Val Asn Thr Gly Gln Ile
227      130      135      140
229 Ser Ile Arg His Gly Met Arg Gly Pro Ser Ser Ala Leu Val Ala Glu
230      145      150      155      160
232 Gln Ala Gly Gly Leu Asp Ala Leu Gly His Ala Arg Arg Thr Ile Arg
233      165      170      175
235 Arg Gly Thr Pro Leu Val Val Ser Gly Gly Val Asp Ser Ala Leu Asp
236      180      185      190
238 Pro Trp Gly Trp Val Ser Glu Ile Ala Ser Gly Arg Ile Ser Thr Ala
239      195      200      205
241 Thr Asp Pro Asp Arg Ala Tyr Leu Pro Phe Asp Glu Arg Ala Ala Gly
242      210      215      220
244 Tyr Val Pro Gly Glu Gly Ala Ile Leu Val Leu Glu Asp Ser Ala
245      225      230      235      240
247 Ala Ala Glu Ala Arg Gly Arg His Asp Ala Tyr Gly Glu Leu Ala Gly
248      245      250      255
250 Cys Ala Ser Thr Phe Asp Pro Ala Pro Gly Ser Gly Arg Pro Ala Gly
251      260      265      270
253 Leu Glu Arg Ala Ile Arg Leu Ala Leu Asn Asp Ala Gly Thr Gly Pro
254      275      280      285
256 Glu Asp Val Asp Val Val Phe Ala Asp Gly Ala Gly Val Pro Glu Leu
257      290      295      300
259 Asp Ala Ala Glu Ala Arg Ala Ile Gly Arg Val Phe Gly Arg Glu Gly
260      305      310      315      320
262 Val Pro Val Thr Val Pro Lys Thr Thr Thr Gly Arg Leu Tyr Ser Gly
263      325      330      335
265 Gly Gly Pro Leu Asp Val Val Thr Ala Leu Met Ser Leu Arg Glu Gly
266      340      345      350
268 Val Ile Ala Pro Thr Ala Gly Val Thr Ser Val Pro Arg Glu Tyr Gly
269      355      360      365
271 Ile Asp Leu Val Leu Gly Glu Pro Arg Ser Thr Ala Pro Arg Thr Ala
272      370      375      380
274 Leu Val Leu Ala Arg Gly Arg Trp Gly Phe Asn Ser Ala Ala Val Leu
275      385      390      395      400
277 Arg Arg Phe Ala Pro Thr Pro
278      405
282 <210> SEQ ID NO: 4
283 <211> LENGTH: 403
284 <212> TYPE: PRT
285 <213> ORGANISM: Saccharopolyspora hirsuta
287 <400> SEQUENCE: 4
288 Met Ser Thr Trp Val Thr Gly Met Gly Val Val Ala Pro Asn Gly Leu
289      1      5      10      15
291 Gly Ala Asp Asp His Trp Ala Ala Thr Leu Lys Gly Arg His Gly Ile
292      20      25      30

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/720,841

DATE: 02/05/2001

TIME: 16:34:36

Input Set : A:\5787585.txt

Output Set: N:\CRF3\02052001\I720841.raw

```

294 Ser Arg Leu Ser Arg Phe Asp Pro Thr Gly Tyr Pro Ala Glu Leu Ala
295      35      40      45
297 Gly Gln Val Leu Asp Phe Asp Ala Thr Glu His Leu Pro Lys Arg Leu
298      50      55      60
300 Leu Pro Gln Thr Asp Val Ser Thr Arg Phe Ala Leu Ala Ala Ala
301 65      70      75      80
303 Trp Ala Leu Ala Asp Ala Glu Val Asp Pro Ala Glu Leu Pro Glu Tyr
304      85      90      95
306 Gly Thr Gly Val Ile Thr Ser Asn Ala Thr Gly Gly Phe Glu Phe Thr
307      100      105      110
309 His Arg Glu Phe Arg Lys Leu Trp Ala Gln Gly Pro Glu Phe Val Ser
310      115      120      125
312 Val Tyr Glu Ser Phe Ala Trp Phe Tyr Ala Val Asn Thr Gly Gln Ile
313      130      135      140
315 Ser Ile Arg His Gly Leu Arg Gly Pro Gly Ser Val Leu Val Ala Glu
316 145      150      155      160
318 Gln Ala Gly Gly Leu Asp Ala Val Gly His Gly Gly Ala Val Arg Asn
319      165      170      175
321 Gly Thr Pro Met Val Val Thr Gly Gly Val Asp Ser Ser Phe Asp Pro
322      180      185      190
324 Trp Gly Trp Val Ser His Val Ser Ser Gly Arg Val Ser Arg Ala Thr
325      195      200      205
327 Asp Pro Gly Arg Ala Tyr Leu Pro Phe Asp Val Ala Ala Asn Gly Tyr
328      210      215      220
330 Val Pro Gly Glu Gly Gly Ala Ile Leu Leu Leu Glu Asp Ala Glu Ser
331 225      230      235      240
333 Ala Lys Ala Arg Gly Ala Thr Gly Tyr Gly Glu Ile Ala Gly Tyr Ala
334      245      250      255
336 Ala Thr Phe Asp Pro Ala Pro Gly Ser Glu Arg Pro Pro Ala Leu Arg
337      260      265      270
339 Arg Ala Ile Glu Leu Ala Leu Ala Asp Ala Glu Leu Arg Pro Glu Gln
340      275      280      285
342 Val Asp Val Val Phe Ala Asp Ala Ala Gly Val Ala Glu Leu Asp Ala
343      290      295      300
345 Ile Glu Ala Ala Ala Ile Arg Glu Leu Phe Gly Pro Ser Gly Val Pro
346 305      310      315      320
348 Val Thr Ala Pro Lys Thr Met Thr Gly Arg Leu Tyr Ser Gly Gly Gly
349      325      330      335
351 Pro Leu Asp Leu Val Ala Ala Leu Leu Ala Ile Arg Asp Gly Val Ile
352      340      345      350
354 Pro Pro Thr Val His Thr Ala Glu Pro Val Pro Glu His Gln Leu Asp
355      355      360      365
357 Leu Val Thr Gly Asp Pro Arg His Gln Gln Leu Gly Thr Ala Leu Val
358      370      375      380
360 Leu Ala Arg Gly Lys Trp Gly Phe Asn Ser Ala Val Val Val Arg Gly
361 385      390      395      400
363 Val Thr Gly
368 <210> SEQ ID NO: 5
369 <211> LENGTH: 415

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/720,841

DATE: 02/05/2001

TIME: 16:34:37

Input Set : A:\5787585.txt

Output Set: N:\CRF3\02052001\I720841.raw

L:14 M:270 C: Current Application Number differs, Replaced Current Application Number

L:15 M:271 C: Current Filing Date differs, Replaced Current Filing Date